



TITLE:

# A singular value decomposition algorithm based on solving hyperplane constrained nonlinear systems

AUTHOR(S):

Yadani, Kenichi; Kondo, Koichi; Iwasaki, Masashi

---

CITATION:

Yadani, Kenichi ...[et al]. A singular value decomposition algorithm based on solving hyperplane constrained nonlinear systems. Applied Mathematics and Computation 2010, 216(3): 779-790

ISSUE DATE:

2010-04-01

URL:

<http://hdl.handle.net/2433/128897>

RIGHT:

© 2010 Elsevier Inc.; この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。; This is not the published version. Please cite only the published version.

# A singular value decomposition algorithm based on solving hyperplane constrained nonlinear systems

Kenichi YADANI<sup>a,\*</sup>, Koichi KONDO<sup>b</sup>, Masashi IWASAKI<sup>c</sup>

<sup>a</sup>*Graduate School of Informatics, Kyoto University, Yoshida-Hommachi, Sakyo-ku, Kyoto, 606-8501, Japan*

<sup>b</sup>*Faculty of Science and Engineering, Doshisha University, 1-3, Tatara Miyakodani, Kyotanabe City, Kyoto, 610-0394, Japan*

<sup>c</sup>*Department of Informatics and Environmental Science, Kyoto Prefectural University, 1-5, Nagaragi-cho, Shimogamo, Sakyo-ku, Kyoto, 606-8522, Japan*

---

## Abstract

A new algorithm for singular value decomposition (SVD) is presented through relating SVD problem to nonlinear systems whose solutions are constrained on hyperplanes. The hyperplane constrained nonlinear systems are solved with the help of Newton's iterative method. It is proved that our SVD algorithm has the quadratic convergence substantially and all singular pairs are computable. These facts are also confirmed by some numerical examples.

*Keywords:* singular value decomposition, Newton's iterative method, nonlinear system, hyperplane, quadratic convergence

---

## 1. Introduction

Singular value decomposition (SVD) is one of important matrix computations in linear algebra. The SVD has so many applications in engineering such as data analysis (cf. [1]), data search (cf. [2]) and image processing (cf. [3]). Almost SVD algorithms, for example the QR (cf. [4]), the qd (cf. [5]) and the dLV algorithms [6, 7], have been designed by the focus on similarity transformation of matrices. In this paper, we propose a new SVD algorithm by relating SVD problem to nonlinear systems.

---

\*Corresponding author

*Email address:* [yadani@amp.i.kyoto-u.ac.jp](mailto:yadani@amp.i.kyoto-u.ac.jp) (Kenichi YADANI)

In [8], Peters and Wilkinson designed an iterative method for computing an eigenpair of matrix based on Newton's iterative method. In [9], Elgindi and Kharab proposed the quadratic method for all eigenpairs with the help of continuation method (cf. [10]). Both methods are derived by replacing eigenvalue problem with nonlinear system. The solutions of the nonlinear system are constrained on a hyperplane. Theoretically, in the iterative method, it is not guaranteed whether all eigenpairs are computed. The quadratic method, however, requires to solve many nonlinear systems associated with the original one. Moreover, the computed eigenpairs are not always with high accuracy. Sometimes it fails to solve eigenpairs. In [11], some of the authors design an algorithm for all eigenpairs through solving the nonlinear systems, which are not equivalent to those in the quadratic method. This algorithm also differs from the quadratic method in that it does not employ the continuation method. The basic idea is to solve the nonlinear systems whose solutions lay on suitable hyperplane with the help of Newton's iterative method sequentially. Generally, it is not easy to find all solutions of nonlinear system by Newton's method. In other words, it is doubtful whether all eigenpairs are computable by only Newton's method. A characteristic parameter associated with hyperplane plays a key role for computing all eigenpairs. The parameter is given from the orthogonal complement of subspace spanned by already obtained eigenvectors. Though some numerical experiments describe that all computed eigenpairs are with high accuracy, the convergence theorem has not been reported yet.

The first purpose of this paper is to design an SVD algorithm along the line similar to [11]. The second is to clarify some properties, especially the convergence, of our SVD algorithm.

This paper is organised as follows. In Section 2, we propose an algorithm for computing a pair of singular value and singular vectors with the help of Newton's method. In Section 3, we design an SVD algorithm based on the algorithm in Section 2. We also discuss the convergence of our SVD algorithm. In Section 4, we investigate some properties concerning the Jacobian matrix appeared in our algorithm. In Section 5, we prove that our SVD algorithm quadratically converges. Numerical examples are demonstrated in Section 6, and in Section 7, conclusion is presented.

## 2. An algorithm to compute a singular pair by solving hyperplane constrained nonlinear system

Elgindi and Kharab [9] first solve an eigenvalue problem by considering a nonlinear system whose solutions are constrained on a hyperplane. The nonlinear system, appeared in [11], is constrained that the solutions are on different hyperplane. With reference to [9] and [11], we begin by deriving a nonlinear system for a singular value and singular vectors.

Let  $A$  be real  $m$ -by- $n$  matrix where  $m \leq n$ . The SVD of  $A$  is given by

$$\begin{aligned} A &= U\Sigma V^T, \quad \Sigma = (S \ O_{m,n-m}) \in \mathbf{R}^{m \times n}, \\ S &= \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m) \in \mathbf{R}^{m \times m}, \end{aligned} \quad (1)$$

where the superscript  $T$  denotes the transposed and  $O_{m,n}$  is  $m$ -by- $n$  zero matrix. Here  $U = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_m) \in \mathbf{R}^{m \times m}$ ,  $V = (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n) \in \mathbf{R}^{n \times n}$  are orthogonal matrices and  $S$  is diagonal matrix with diagonal entries  $\sigma_1, \sigma_2, \dots, \sigma_m$ . In the case where  $m > n$ , it is easy to discuss the SVD by considering  $A^T$ . Each diagonal entry  $\sigma_k$  in  $S$  is the singular value of  $A$  and the vectors  $\mathbf{u}_k, \mathbf{v}_k$  are the singular vectors corresponding to  $\sigma_k$ . The vectors  $\mathbf{u}_k, \mathbf{v}_k$  are called the left, the right singular vectors, respectively. For simplicity, let us call  $(\sigma_k, \mathbf{u}_k, \mathbf{v}_k)$  the singular pair.

Equation (1) is also expressed as

$$\begin{aligned} \begin{cases} A\mathbf{v}_k = \sigma_k \mathbf{u}_k, & A^T \mathbf{u}_k = \sigma_k \mathbf{v}_k, & k = 1, 2, \dots, m, \\ \|\mathbf{u}_k\|_2 = 1, & \|\mathbf{v}_k\|_2 = 1, \end{cases} \quad (2) \\ A\mathbf{v}_k = \mathbf{0}, \quad \|\mathbf{v}_k\|_2 = 1, \quad k = m+1, m+2, \dots, n. \quad (3) \end{aligned}$$

All singular pairs  $(\sigma_k, \mathbf{u}_k, \mathbf{v}_k)$  for  $k = 1, 2, \dots, m$  and the right singular vectors  $\mathbf{v}_k$  for  $k = m+1, m+2, \dots, n$  are computable by solving (2) and (3), respectively. The solutions of (3) are given by the orthonormal bases of  $\mathcal{V}_m^\perp$ , where  $\mathcal{V}_k = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  is a real linear space spanned by  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ , and  $\mathcal{V}_k^\perp$  is the orthogonal complement of  $\mathcal{V}_k$ . So, it is substantially necessary to discuss how to compute only the singular pairs  $(\sigma_k, \mathbf{u}_k, \mathbf{v}_k)$  for  $k = 1, 2, \dots, m$ . It is obvious that, for real  $\alpha$ ,  $\alpha \mathbf{u}_k$  and  $\alpha \mathbf{v}_k$  are also the solution of the 1st and the 2nd equations of (2). The 3rd and the 4th of (2) are for the normalizations of  $\mathbf{u}_k$  and  $\mathbf{v}_k$ , respectively. Thus, for the pair  $(\sigma_k, \alpha_k \mathbf{u}_k, \alpha_k \mathbf{v}_k)$ , we may impose the different constriction from the normalizations  $\|\mathbf{u}_k\|_2 = 1$  and  $\|\mathbf{v}_k\|_2 = 1$ . Let  $\mathbf{u}$  be on the hyperplane

$\Gamma = \{\mathbf{u} \mid \mathbf{u} \in \mathbf{R}^m, (\mathbf{z}, \mathbf{u}) = C\}$  where  $\mathbf{z} \in \mathbf{R}^m$  is a normal vector and  $C \neq 0$  is a constant. Then, instead of (2), we derive

$$A\mathbf{v} = \sigma\mathbf{u}, \quad A^T\mathbf{u} = \sigma\mathbf{v}, \quad (\mathbf{z}, \mathbf{u}) = C, \quad (4)$$

where  $(\cdot, \cdot)$  denotes inner product. The above discussion gives rise to the following theorem for the solutions of (4).

**Theorem 2.1.** *If  $\mathbf{z}$  satisfies  $(\mathbf{z}, \mathbf{u}_k) \neq 0$  for  $k = 1, 2, \dots, m$ , then the solutions of (4) are*

$$(\sigma, \mathbf{u}, \mathbf{v}) = (\sigma_k, \alpha_k \mathbf{u}_k, \alpha_k \mathbf{v}_k), \quad \alpha_k = \frac{C}{(\mathbf{z}, \mathbf{u}_k)}, \quad k = 1, 2, \dots, m. \quad (5)$$

The 1st and the 3rd equations of (4) yield that  $\sigma = (A^T \mathbf{z}, \mathbf{v})/C$ . Here we regard  $\sigma$  as the function of  $\mathbf{v}$ , namely,

$$\sigma(\mathbf{v}) = \frac{(\mathbf{w}, \mathbf{v})}{C}, \quad \mathbf{w} = A^T \mathbf{z}. \quad (6)$$

Moreover, (4) is rewritten as the quadratic nonlinear system

$$\mathbf{H}(\mathbf{x}) = \mathbf{0}, \quad (7)$$

where

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} A\mathbf{v} - \sigma(\mathbf{v})\mathbf{u} \\ A^T\mathbf{u} - \sigma(\mathbf{v})\mathbf{v} \end{pmatrix}, \quad (8)$$

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \in \mathbf{R}^{m+n}. \quad (9)$$

It is remarkable here that the solutions of (7) are equal to those of (4). By combining it with Theorem 2.1, the solutions of (7) become (5). Thus, the singular pairs are computable through solving nonlinear system (7) with the help of Newton's iterative method. Let us define

$$\mathbf{x}^{(\ell)} = \begin{pmatrix} \mathbf{u}^{(\ell)} \\ \mathbf{v}^{(\ell)} \end{pmatrix},$$

where  $\mathbf{u}^{(\ell)}$ ,  $\mathbf{v}^{(\ell)}$  and  $\mathbf{x}^{(\ell)}$  denote the values of  $\mathbf{u}$ ,  $\mathbf{v}$  and  $\mathbf{x}$  at the iteration number  $\ell$ , respectively. The sequence  $\{\mathbf{x}^{(\ell)}\}_{\ell=0,1,\dots}$ , whose limit becomes the

solution of (7) as  $\ell \rightarrow \infty$  if converges, is given by

$$\mathbf{x}^{(\ell+1)} = \Phi(\mathbf{x}^{(\ell)}), \quad (10)$$

$$\Phi(\mathbf{x}) = \mathbf{x} - (J(\mathbf{x}))^{-1} \mathbf{H}(\mathbf{x}), \quad (11)$$

$$J(\mathbf{x}) = \begin{pmatrix} -\sigma(\mathbf{v})I_m & A - \mathbf{u}\mathbf{w}^T C^{-1} \\ A^T & -\sigma(\mathbf{v})I_n - \mathbf{v}\mathbf{w}^T C^{-1} \end{pmatrix}. \quad (12)$$

Here  $J(\mathbf{x})$  is the Jacobian matrix of  $\mathbf{H}(\mathbf{x})$  and  $I_k$  is  $k$ -by- $k$  identity matrix. We obtain a remarkable property for the existence range of  $\mathbf{u}^{(\ell)}$  in  $\mathbf{x}^{(\ell)}$  as below.

**Lemma 2.2.** *Let  $\mathbf{u}^{(\ell)}$  in  $\mathbf{x}^{(\ell)}$  satisfy  $(\mathbf{z}, \mathbf{u}^{(\ell)}) = C$  at some  $\ell$ . If  $J(\mathbf{x}^{(\ell)})$  is nonsingular, then  $\mathbf{u}^{(\ell+1)}$  in  $\mathbf{x}^{(\ell+1)}$ , generated by (10), satisfies  $(\mathbf{z}, \mathbf{u}^{(\ell+1)}) = C$ . Namely,  $\mathbf{u}^{(\ell+1)}$  is also on the hyperplane  $\Gamma$  if  $\mathbf{u}^{(\ell)}$  is on  $\Gamma$ .*

PROOF. From (10) and (11), it follows that

$$J(\mathbf{x}^{(\ell)})(\mathbf{x}^{(\ell+1)} - \mathbf{x}^{(\ell)}) = -\mathbf{H}(\mathbf{x}^{(\ell)}). \quad (13)$$

Substituting (8) and (12) into (13), we have

$$-\sigma(\mathbf{v}^{(\ell)})\mathbf{u}^{(\ell+1)} + (A - \mathbf{u}^{(\ell)}\mathbf{w}^T C^{-1})\mathbf{v}^{(\ell+1)} = -\sigma(\mathbf{v}^{(\ell)})\mathbf{u}^{(\ell)}, \quad (14)$$

$$A^T \mathbf{u}^{(\ell+1)} - (\sigma(\mathbf{v}^{(\ell)})I_n + \mathbf{v}^{(\ell)}\mathbf{w}^T C^{-1})\mathbf{v}^{(\ell+1)} = -\sigma(\mathbf{v}^{(\ell)})\mathbf{v}^{(\ell)}. \quad (15)$$

Equations (14) and (15) lead to

$$\mathbf{v}^{(\ell+1)} = -\sigma(\mathbf{v}^{(\ell)})(M + \mathbf{p}\mathbf{q}^T)^{-1}\mathbf{p},$$

where  $M = A^T A - \sigma(\mathbf{v}^{(\ell)})^2 I_n$ ,  $\mathbf{p} = A^T \mathbf{u}^{(\ell)} + \sigma(\mathbf{v}^{(\ell)})\mathbf{v}^{(\ell)}$  and  $\mathbf{q} = -\mathbf{w}C^{-1}$ . Applying the Sherman-Morrison formula (cf. [12]), namely,

$$(M + \mathbf{p}\mathbf{q}^T)^{-1} = \left( I_n - \frac{M^{-1}\mathbf{p}\mathbf{q}^T}{1 + (\mathbf{q}, M^{-1}\mathbf{p})} \right) M^{-1},$$

it follows that

$$\mathbf{u}^{(\ell+1)} = \frac{\mathbf{u}^{(\ell)} - A\bar{\mathbf{v}}^{(\ell)}}{1 - (\mathbf{w}, \bar{\mathbf{v}}^{(\ell)})C^{-1}}, \quad \mathbf{v}^{(\ell+1)} = -\frac{\sigma(\mathbf{v}^{(\ell)})\bar{\mathbf{v}}^{(\ell)}}{1 - (\mathbf{w}, \bar{\mathbf{v}}^{(\ell)})C^{-1}}, \quad (16)$$

where  $\bar{\mathbf{v}}^{(\ell)} = (A^T A - \sigma(\mathbf{v}^{(\ell)})^2 I_n)^{-1}(A^T \mathbf{u}^{(\ell)} + \sigma(\mathbf{v}^{(\ell)})\mathbf{v}^{(\ell)})$ . It is of significance to note here that (16) is well-defined if  $J(\mathbf{x}^{(\ell)})$  is nonsingular. See Section 4

for the nonsingularity of  $J(\mathbf{x}^{(\ell)})$ . Thus, by using  $\mathbf{w} = A^T \mathbf{z}$  and  $(\mathbf{z}, \mathbf{u}^{(\ell)}) = C$ , from the 1st equation of (16), we derive

$$(\mathbf{z}, \mathbf{u}^{(\ell+1)}) = \frac{(\mathbf{z}, \mathbf{u}^{(\ell)}) - (\mathbf{z}, A\bar{\mathbf{v}}^{(\ell)})}{1 - (\mathbf{w}, \bar{\mathbf{v}}^{(\ell)})C^{-1}} = \frac{C - (\mathbf{w}, \bar{\mathbf{v}}^{(\ell)})}{1 - (\mathbf{w}, \bar{\mathbf{v}}^{(\ell)})C^{-1}} = C.$$

Therefore, it is concluded that  $\mathbf{u}^{(\ell+1)}$  is also on the hyperplane  $\Gamma$  with  $\mathbf{u}^{(\ell)}$ .  
□

The following theorem is given by using Lemma 2.2 repeatedly.

**Theorem 2.3.** *Let  $\mathbf{u}^{(0)}$  in  $\mathbf{x}^{(0)}$  be on the hyperplane  $\Gamma$ . If  $J(\mathbf{x}^{(\ell)})$  is nonsingular, then  $\mathbf{u}^{(\ell)}$  in  $\mathbf{x}^{(\ell)}$ , generated by (10), are on  $\Gamma$  for  $\ell = 0, 1, \dots$*

Theorem 2.3 claims that  $\mathbf{u}^{(\ell)}$  in  $\mathbf{x}^{(\ell)}$  for  $\ell = 0, 1, \dots$  are theoretically on the hyperplane  $\Gamma$  with  $\mathbf{u}^{(0)}$  in  $\mathbf{x}^{(0)}$ . In finite arithmetic,  $\mathbf{u}^{(\ell)}$  may not be exactly computed. In other words,  $\mathbf{u}^{(\ell+1)}$ , generated by (10), is not always on  $\Gamma$ . So, for solving (7), we adopt the mapping from  $\mathbf{x}^{(\ell)}$  to  $\mathbf{x}^{(\ell+1)}$  given as

$$\mathbf{x}^{(\ell+1)} = \Psi(\Phi(\mathbf{x}^{(\ell)})), \quad \ell = 0, 1, \dots, \quad (17)$$

$$\Psi(\mathbf{x}) = \frac{C}{(\mathbf{z}, \mathbf{u})} \mathbf{x}. \quad (18)$$

Here  $\Phi$  is the same as in (11). By the mapping  $\Psi$ ,  $\mathbf{u}^{(\ell+1)}$  in  $\mathbf{x}^{(\ell+1)}$  is surely on the hyperplane  $\Gamma$ . Of course,  $\mathbf{u}^{(0)}$  in  $\mathbf{x}^{(0)}$  should be selected as  $(\mathbf{z}, \mathbf{u}^{(0)}) = C$ . If  $(\mathbf{z}, \mathbf{u}^{(0)}) \neq C$ , then we replace  $\mathbf{x}^{(0)}$  with  $C\mathbf{x}^{(0)}/(\mathbf{z}, \mathbf{u}^{(0)})$ . The above discussion leads to the theorem similar to Theorem 2.3.

**Theorem 2.4.** *Let  $\mathbf{u}^{(0)}$  in  $\mathbf{x}^{(0)}$  be on the hyperplane  $\Gamma$ . If  $J(\mathbf{x}^{(\ell)})$  is nonsingular, then  $\mathbf{u}^{(\ell)}$  in  $\mathbf{x}^{(\ell)}$ , generated by (17), are on  $\Gamma$  for  $\ell = 0, 1, \dots$*

PROOF. Obviously, from (17), (18) and Theorem 2.3. □

Suppose that  $\mathbf{H}(\mathbf{x}^{(\ell)})$  converges to  $\mathbf{0}$  as  $\ell \rightarrow \infty$ . Then  $\mathbf{x}^{(\ell)}$  converges to one of the solutions (5). Namely,  $\mathbf{x}^{(\ell)} \rightarrow \alpha_k \mathbf{x}_k$ , where

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{u}_k \\ \mathbf{v}_k \end{pmatrix}.$$

At  $\ell = \ell^*$ , if  $\mathbf{x}^{(\ell^*)}$  satisfies

$$\|\mathbf{H}(\mathbf{x}^{(\ell^*)})\|_\infty < \varepsilon_{\text{tol}} \|\mathbf{u}^{(\ell^*)}\|_\infty \quad (19)$$

for small positive  $\varepsilon_{\text{tol}}$ , then  $(\sigma(\mathbf{v}^{(\ell^*)}), \mathbf{u}^{(\ell^*)}, \mathbf{v}^{(\ell^*)})$  becomes an approximation of  $(\sigma_k, \alpha_k \mathbf{u}_k, \alpha_k \mathbf{v}_k)$ . Especially, in finite arithmetic, the mapping (17) stops at  $\ell = \ell_{\text{max}}$  even if (19) does not hold. The variable  $\sigma(\mathbf{v}^{(\ell)})$  sometimes becomes negative in (17). In that case, we may replace  $\mathbf{v}^{(\ell+1)}$  with  $-\mathbf{v}^{(\ell+1)}$  for keeping  $\sigma(\mathbf{v}^{(\ell+1)}) \geq 0$ .

The algorithm for a singular pair, by (6), (8), (9), (11), (12) and (17)–(19), is summarised as follows.

**Algorithm 1.**

**Input:**  $A \in \mathbf{R}^{m \times n}$ ,  $\mathbf{u}^{(0)} \in \mathbf{R}^m$ ,  $\mathbf{v}^{(0)} \in \mathbf{R}^n$ ,  $\mathbf{z} \in \mathbf{R}^m$ ,  $C \in \mathbf{R}$  ( $C \neq 0$ ),  
 $\varepsilon_{\text{tol}} > 0$ ,  $\ell_{\text{max}} > 0$ .

**Output:**  $\sigma^* \geq 0$ ,  $\mathbf{u}^* \in \mathbf{R}^m$ ,  $\mathbf{v}^* \in \mathbf{R}^n$ .

01.  $\mathbf{w} := A^T \mathbf{z}$ ,  $\mathbf{u}^{(0)} := C \mathbf{u}^{(0)} / (\mathbf{z}, \mathbf{u}^{(0)})$ ,  $\mathbf{v}^{(0)} := C \mathbf{v}^{(0)} / (\mathbf{z}, \mathbf{u}^{(0)})$ ,  
 $\sigma^{(0)} := (\mathbf{w}, \mathbf{v}^{(0)}) / C$ .
02. **if**  $\sigma^{(0)} < 0$  **then**  $\sigma^{(0)} := -\sigma^{(0)}$ ,  $\mathbf{v}^{(0)} := -\mathbf{v}^{(0)}$ .
03.  $\mathbf{H}^{(0)} := \begin{pmatrix} A \mathbf{v}^{(0)} - \sigma^{(0)} \mathbf{u}^{(0)} \\ A^T \mathbf{u}^{(0)} - \sigma^{(0)} \mathbf{v}^{(0)} \end{pmatrix}$ ,  $\ell := 0$ .
04. **while**  $(\|\mathbf{H}^{(\ell)}\|_\infty \geq \varepsilon_{\text{tol}} \|\mathbf{u}^{(\ell)}\|_\infty \text{ and } \ell \leq \ell_{\text{max}})$
05.  $J^{(\ell)} := \begin{pmatrix} -\sigma^{(\ell)} I_m & A - \mathbf{u}^{(\ell)} \mathbf{w}^T C^{-1} \\ A^T & -\sigma^{(\ell)} I_n - \mathbf{v}^{(\ell)} \mathbf{w}^T C^{-1} \end{pmatrix}$ .
06.  $\begin{pmatrix} \hat{\mathbf{u}}^{(\ell+1)} \\ \hat{\mathbf{v}}^{(\ell+1)} \end{pmatrix} := \begin{pmatrix} \mathbf{u}^{(\ell)} \\ \mathbf{v}^{(\ell)} \end{pmatrix} - (J^{(\ell)})^{-1} \mathbf{H}^{(\ell)}$ .
07.  $\mathbf{u}^{(\ell+1)} := C \hat{\mathbf{u}}^{(\ell+1)} / (\mathbf{z}, \hat{\mathbf{u}}^{(\ell+1)})$ ,  $\mathbf{v}^{(\ell+1)} := C \hat{\mathbf{v}}^{(\ell+1)} / (\mathbf{z}, \hat{\mathbf{u}}^{(\ell+1)})$ .
08.  $\sigma^{(\ell+1)} := (\mathbf{w}, \mathbf{v}^{(\ell+1)}) / C$ .
09. **if**  $\sigma^{(\ell+1)} < 0$  **then**  $\sigma^{(\ell+1)} := -\sigma^{(\ell+1)}$ ,  $\mathbf{v}^{(\ell+1)} := -\mathbf{v}^{(\ell+1)}$ .
10.  $\mathbf{H}^{(\ell+1)} := \begin{pmatrix} A \mathbf{v}^{(\ell+1)} - \sigma^{(\ell+1)} \mathbf{u}^{(\ell+1)} \\ A^T \mathbf{u}^{(\ell+1)} - \sigma^{(\ell+1)} \mathbf{v}^{(\ell+1)} \end{pmatrix}$ .
11.  $\ell := \ell + 1$ .
12. **end**
13. **return**  $(\sigma^*, \mathbf{u}^*, \mathbf{v}^*) := (\sigma^{(\ell)}, \mathbf{u}^{(\ell)}, \mathbf{v}^{(\ell)})$

The pair  $(\sigma^*, \mathbf{u}^*, \mathbf{v}^*)$  in the 13th line of Algorithm 1 also becomes the solutions of (4). Namely, the pair  $(\sigma_k, \alpha_k \mathbf{u}_k, \alpha_k \mathbf{v}_k)$  are computed by Algorithm 1. The existence conditions of  $(J(\mathbf{x}^{(\ell)}))^{-1}$  in the 6th line are clarified in Section 4. Especially, we call the iteration from the 4th to the 12th Newton type iteration as below. The iteration number of Newton type iteration is counted on  $\ell$ , and then Newton type iteration coercively stops if  $\ell$  becomes



larger than some integer  $\ell_{\max}$ .

### 3. An SVD algorithm based on solving hyperplane constrained nonlinear systems

In this section, we propose an SVD algorithm for computing not only a singular pair but also all singular pairs of rectangular matrix  $A$  by using Algorithm 1 sequentially.

It is remarkable that the computable pair  $(\sigma^*, \mathbf{u}^*, \mathbf{v}^*)$  in the 13th line of Algorithm 1 is changeable by the initial  $\mathbf{u}^{(0)}, \mathbf{v}^{(0)}$ . Theoretically, the other pairs are computed by changing  $\mathbf{u}^{(0)}, \mathbf{v}^{(0)}$ . All singular pairs are accordingly computed by using Algorithm 1 repeatedly. However, in Newton's iterative method, the mapping from initial values to their limits generates fractal graph (cf. [13]). Thus, it is not easy to select an initial value for a desired singular pair. One of the settlement strategies is to introduce the continuation method (cf. [10]). In [9], Elgindi and Kharab proposed the quadratic method with the continuation method for all eigenpairs. By some numerical experiments, it is observed that the quadratic method does not always converge. Sometimes the computed eigenpairs are not with high accuracy even if the quadratic method converges. It also seems to be not easy to apply the continuation method as in [9] to the SVD problems. On the other hand, in [11], some of the authors designed an algorithm, without the continuation method, through solving nonlinear systems sequentially. In this section, we derive a new algorithm for SVD along the line similar to [11].

Let us begin by giving a lemma for the solutions of (4).

**Lemma 3.1.** *If  $\mathbf{z}$  satisfies  $(\mathbf{z}, \mathbf{u}_k) = 0$ , then the solutions of (4) are*

$$(\sigma, \mathbf{u}, \mathbf{v}) = (\sigma_i, \alpha_i \mathbf{u}_i, \alpha_i \mathbf{v}_i), \quad i = 1, \dots, k-1, k+1, \dots, m.$$

PROOF. Let us recall that  $C \neq 0$  in (4). Since  $(\mathbf{z}, \mathbf{u}_k) = 0$ , it is obvious that  $\mathbf{u} = \mathbf{u}_k$  is not a solution of (4). And then  $\mathbf{u} = \mathbf{u}_i$  for  $i \neq k$  is the solution of (4).  $\square$

Let  $\mathcal{U}_k$  denotes  $\mathcal{U}_k = \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ . Then the following theorem holds immediately.

**Theorem 3.2.** *Let  $\mathcal{U}_k^\perp$  be the orthogonal complement of subspace  $\mathcal{U}_k$ . If  $\mathbf{z} \in \mathcal{U}_k^\perp$ , then the solutions of (4) are*

$$(\sigma, \mathbf{u}, \mathbf{v}) = (\sigma_i, \alpha_i \mathbf{u}_i, \alpha_i \mathbf{v}_i), \quad i = k+1, k+2, \dots, m. \quad (20)$$

Theorem 3.2 implies that the solutions (20) exist in  $\mathcal{U}_k^\perp$ . Suppose that  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  are already obtained by Algorithm 1, and normal vector  $\mathbf{z}$  of  $\Gamma$  is given as  $\mathbf{z} \in \mathcal{U}_k^\perp$ . Then we obtain the following theorem.

**Theorem 3.3.** *Suppose that  $\mathbf{z} \in \mathcal{U}_k^\perp$ . If  $\mathbf{x}^{(\ell)}$ , generated by (17), converges to*

$$\mathbf{x}^* = \begin{pmatrix} \mathbf{u}^* \\ \mathbf{v}^* \end{pmatrix}$$

*as  $\ell \rightarrow \infty$ , then  $\mathbf{u}^*$  satisfies  $\mathbf{u}^* \notin \mathcal{U}_k$ .*

PROOF. Obviously,  $\Gamma \cap \mathcal{U}_k = \emptyset$ . Let us recall that  $\mathbf{u}^{(\ell)} \in \Gamma$  for  $\ell = 0, 1, \dots$  shown in Theorem 2.4. If  $\mathbf{x}^{(\ell)}$  converges to  $\mathbf{x}^*$ , then  $\mathbf{u}^* \in \Gamma$ , accordingly,  $\mathbf{u}^* \notin \mathcal{U}_k$ .  $\square$

From Theorem 3.3, it turns out that  $\mathbf{u}^*/\|\mathbf{u}^*\|_2$  is not equal to  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ . Thus, the pair  $(\sigma_{k+1}, \mathbf{u}_{k+1}, \mathbf{v}_{k+1}) = (\sigma(\mathbf{v}^*), \mathbf{u}^*/\|\mathbf{u}^*\|_2, \mathbf{v}^*/\|\mathbf{v}^*\|_2)$  is different from the already obtained singular pairs. All singular pairs are computable by selecting  $\mathbf{z}$  from the orthogonal complement  $\mathcal{U}_k^\perp$  and using Algorithm 1 repeatedly. Therefore, a new SVD algorithm is given as follows.

### Algorithm 2.

**Input:**  $A \in \mathbf{R}^{m \times n}$ ,  $\varepsilon > 0$ ,  $C \in \mathbf{R}$  ( $C \neq 0$ ),  $\varepsilon_{\text{tol}} > 0$ ,  $\ell_{\text{max}} > 0$ .

**Output:**  $U = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_m) \in \mathbf{R}^{m \times m}$ ,  $V = (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n) \in \mathbf{R}^{n \times n}$ ,  
 $\Sigma = (\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m) \ O_{n-m, n})$ .

01. **for**  $k = 1, 2, \dots, m$
02. Randomly select  $\mathbf{z} \in \mathbf{R}^m$ ,  $\mathbf{u}^{(0)} \in \mathbf{R}^m$ ,  $\mathbf{v}^{(0)} \in \mathbf{R}^n$ .
03. Orthonormalize  $\mathbf{z}$  to  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}\}$  by the Gram-Schmidt process.
04. Orthonormalize  $\mathbf{u}^{(0)}$  to  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}\}$  by the Gram-Schmidt process.
05. Orthonormalize  $\mathbf{v}^{(0)}$  to  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}\}$  by the Gram-Schmidt process.
06. Compute  $(\sigma^*, \mathbf{u}^*, \mathbf{v}^*)$  by Algorithm 1 with inputs  $A$ ,  $\mathbf{u}^{(0)}$ ,  $\mathbf{v}^{(0)}$ ,  $\mathbf{z}$ ,  $C$ ,  $\varepsilon_{\text{tol}}$ ,  $\ell_{\text{max}}$ .
07.  $\sigma_k := \sigma^*$ ,  $\mathbf{u}_k := \mathbf{u}^*/\|\mathbf{u}^*\|_2$ ,  $\mathbf{v}_k := \mathbf{v}^*/\|\mathbf{v}^*\|_2$ .
08. **if**  $(\|A\mathbf{v}_k - \sigma_k \mathbf{u}_k\|_\infty \geq \varepsilon \text{ or } \|A^T \mathbf{u}_k - \sigma_k \mathbf{v}_k\|_\infty \geq \varepsilon)$  **then goto** 02.
09. **end**
10. **for**  $k = m + 1, m + 2, \dots, n$
11. Randomly select  $\mathbf{v}_k \in \mathbf{R}^n$ .

12. Orthonormalize  $\mathbf{v}_k$  to  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}\}$  by the Gram-Schmidt process.
13. **end**
14. **return**  $U := (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_m), V := (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n),$   
 $\Sigma := (\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m) \ O_{n-m,n}).$

In the 2nd line,  $\mathbf{z}$  is randomly given as  $\mathbf{z} \in \mathbf{R}^m$ . The initial values  $\mathbf{u}^{(0)}$  and  $\mathbf{v}^{(0)}$  in Algorithm 1 are also randomly given as  $\mathbf{u}^{(0)} \in \mathbf{R}^m$  and  $\mathbf{v}^{(0)} \in \mathbf{R}^n$ , respectively. In the 3rd,  $\mathbf{z}$  is orthonormalized to  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$ , namely,  $(\mathbf{z}, \mathbf{u}_i) = 0$  for  $i = 1, 2, \dots, k-1$ , by the Gram-Schmidt process. Similarly, in the 4th and the 5th,  $\mathbf{u}^{(0)}$  and  $\mathbf{v}^{(0)}$  are given as  $(\mathbf{u}^{(0)}, \mathbf{u}_i) = 0$  and  $(\mathbf{v}^{(0)}, \mathbf{v}_i) = 0$  for  $i = 1, 2, \dots, k-1$ , respectively. This is because that  $\mathbf{z} \in \mathcal{U}_{k-1}^\perp$ ,  $\mathbf{u}_k \in \mathcal{U}_{k-1}^\perp$  and  $\mathbf{v}_k \in \mathcal{V}_{k-1}^\perp$ . In the 6th,  $(\sigma^*, \mathbf{u}^*, \mathbf{v}^*)$  is computed by Algorithm 1. A singular pair  $(\sigma_k, \mathbf{u}_k, \mathbf{v}_k)$  is given by the normalization of  $(\sigma^*, \mathbf{u}^*, \mathbf{v}^*)$  in the 7th. In the 8th, we regard that Algorithm 1 does not converge, if the residual norms  $\|A\mathbf{v}_k - \sigma_k \mathbf{u}_k\|_\infty$  and  $\|A^T \mathbf{u}_k - \sigma_k \mathbf{v}_k\|_\infty$  are not smaller than  $\varepsilon$ . And then, we return to the 2nd. We try to recompute by changing  $\mathbf{u}^{(0)}$ ,  $\mathbf{v}^{(0)}$  and  $\mathbf{z}$ . From the 10th to the 13th, the solutions of (3) are computed from  $\mathbf{v}_k \in \mathcal{V}_{k-1}^\perp$  for  $k = m+1, m+2, \dots, n$ .

From Theorem 3.3, we derive the following theorem for Algorithm 2 immediately.

**Theorem 3.4.** *All of singular pairs of  $A$  are computable if Algorithm 1 always converges in Algorithm 2.*

Theorem 3.4 shows that the convergence of Algorithm 2 is dependent on that of Newton type iteration in Algorithm 1. The convergence condition of Newton type iteration is discussed in the following sections.

#### 4. Nonsingularity and boundedness of Jacobian matrix

In order to analyse the convergence of Newton type iteration in the later section, it is necessary to prove that the Jacobian matrix  $J(\mathbf{x})$  is nonsingular and bounded. In this section, we clarify the nonsingularity of  $J(\mathbf{x})$  in the both cases  $\mathbf{x} = \alpha_k \mathbf{x}_k$  and  $\mathbf{x} \neq \alpha_k \mathbf{x}_k$ . Simultaneously, the boundedness of  $J(\mathbf{x})$  is also shown.

Note that  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ ,  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  are orthonormal bases of  $\mathbf{R}^m$ ,  $\mathbf{R}^n$ , respectively. Then  $\mathbf{u} \in \mathbf{R}^m$ ,  $\mathbf{v} \in \mathbf{R}^n$  and  $\mathbf{z} \in \mathbf{R}^m$  are uniquely given as

$$\mathbf{u} = \sum_{i=1}^m \tilde{u}_i \mathbf{u}_i, \quad \mathbf{v} = \sum_{i=1}^n \tilde{v}_i \mathbf{v}_i, \quad \mathbf{z} = \sum_{i=1}^m \tilde{z}_i \mathbf{u}_i. \quad (21)$$

Of course,  $\tilde{u}_i = (\mathbf{u}_i, \mathbf{u})$ ,  $\tilde{z}_i = (\mathbf{u}_i, \mathbf{z})$  for  $i = 1, 2, \dots, m$  and  $\tilde{v}_i = (\mathbf{v}_i, \mathbf{v})$  for  $i = 1, 2, \dots, n$ . With those notation, we give a lemma for the Jacobian  $\det(J(\mathbf{x}))$ .

**Lemma 4.1.** *The Jacobian  $\det(J(\mathbf{x}))$  has the following expression*

$$\det(J(\mathbf{x})) = (-1)^n \sigma(\mathbf{v})^{n-m} \times \det \left( \left( (\sigma_i^2 - \sigma(\mathbf{v})^2) \delta_{ij} - \frac{\sigma_i \tilde{u}_i + \sigma(\mathbf{v}) \tilde{v}_i}{C} \sigma_j \tilde{z}_j \right)_{i,j=1,1}^{m,m} \right), \quad (22)$$

where  $\delta_{ij}$  is the Kronecker's delta.

PROOF. It is shown in [12] that, for 2-by-2 block matrix,

$$\det \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \det(P) \det(S - RP^{-1}Q), \quad (23)$$

where  $P$  is a nonsingular square matrix. Let  $P = -\sigma(\mathbf{v})I_m$ ,  $Q = A - \mathbf{u}\mathbf{w}^T C^{-1}$ ,  $R = A^T$  and  $S = -\sigma(\mathbf{v})I_n - \mathbf{v}\mathbf{w}^T C^{-1}$  in (23). Comparing it with (12), we derive

$$\begin{aligned} \det(J(\mathbf{x})) &= (-1)^m \sigma(\mathbf{v})^{m-n} \det \left( (A^T A - \sigma(\mathbf{v})^2 I_n) - (A^T \mathbf{u} + \sigma(\mathbf{v}) \mathbf{v}) \mathbf{w}^T C^{-1} \right). \end{aligned}$$

Let  $\tilde{u}_i, \tilde{v}_i$  and  $\tilde{z}_i$  be the  $i$ th entries in  $\tilde{\mathbf{u}}, \tilde{\mathbf{v}}$  and  $\tilde{\mathbf{z}}$ , respectively. Then, from (21),  $\mathbf{u} = U\tilde{\mathbf{u}}$ ,  $\mathbf{v} = V\tilde{\mathbf{v}}$  and  $\mathbf{z} = U\tilde{\mathbf{z}}$ . Let us recall that  $A = U\Sigma V^T$  and  $\mathbf{w} = A^T \mathbf{z}$ . Then  $A^T A = V\Sigma^T \Sigma V^T$ ,  $\mathbf{w}^T = \tilde{\mathbf{z}}^T \Sigma V^T$  and  $A^T \mathbf{u} = V\Sigma^T \tilde{\mathbf{u}}$ . By taking account that  $\det(V) \det(V^T) = 1$ , it follows that

$$\begin{aligned} \det(J(\mathbf{x})) &= (-1)^m \sigma(\mathbf{v})^{m-n} \\ &\times \det \left( (\Sigma^T \Sigma - \sigma(\mathbf{v})^2 I_n) - (\Sigma^T \tilde{\mathbf{u}} + \sigma(\mathbf{v}) \tilde{\mathbf{v}}) \tilde{\mathbf{z}}^T \Sigma C^{-1} \right). \quad (24) \end{aligned}$$

Moreover, by adjusting the determinant in the right-hand side of (24),

$$\det(J(\mathbf{x})) = (-1)^m \sigma(\mathbf{v})^{m-n} \det \begin{pmatrix} \tilde{P} & O_{m,n-m} \\ * & -\sigma(\mathbf{v})^2 I_{n-m} \end{pmatrix}, \quad (25)$$

where the  $(i, j)$  entry in  $\tilde{P}$  is given by  $\tilde{P}_{ij} = (\sigma_i^2 - \sigma(\mathbf{v})^2) \delta_{ij} - (\sigma_i \tilde{u}_i + \sigma(\mathbf{v}) \tilde{v}_i) \sigma_j \tilde{z}_j C^{-1}$ . Substituting  $\det(-\sigma(\mathbf{v})^2 I_{n-m}) = (-1)^{n-m} \sigma(\mathbf{v})^{2(n-m)}$  into (25), the Jacobian  $\det(J(\mathbf{x}))$  becomes  $(-1)^n \sigma(\mathbf{v})^{n-m} \det(\tilde{P})$ . Namely, we have (22).  $\square$

Lemma 4.1 leads to the next theorems for the nonsingularity of the Jacobian matrix  $J(\mathbf{x})$ .

**Theorem 4.2.** *Suppose that  $\mathbf{z}$  satisfies  $(\mathbf{z}, \mathbf{u}_k) \neq 0$ . If the singular values of  $A$  are distinct and not zero, then the Jacobian matrix  $J(\mathbf{x})$  is nonsingular at  $\mathbf{x} = \alpha_k \mathbf{x}_k$ .*

PROOF. Let  $\mathbf{u} = \alpha_k \mathbf{u}_k$ ,  $\mathbf{v} = \alpha_k \mathbf{v}_k$  in (22). Noting that  $\sigma(\alpha_k \mathbf{v}_k) = \sigma_k$ ,  $\tilde{u}_i = \alpha_k \delta_{ik}$ ,  $\tilde{v}_i = \alpha_k \delta_{ik}$  and  $\alpha_k = C/(\mathbf{z}, \mathbf{u}_k) = C/\tilde{z}_k$ , we have

$$\begin{aligned} \det(J(\alpha_k \mathbf{x}_k)) &= (-1)^n \sigma_k^{n-m} \det \left( \left( (\sigma_i^2 - \sigma_k^2) \delta_{ij} - \frac{2\sigma_k \alpha_k \delta_{ik} \sigma_j \tilde{z}_j}{C} \right)_{i,j=1,1}^{m,m} \right) \\ &= \frac{2(-1)^{n+1} \alpha_k \sigma_k^{n-m+1}}{C} \\ &\quad \times \det \begin{pmatrix} \sigma_1^2 - \sigma_k^2 & & & & & \\ & \sigma_2^2 - \sigma_k^2 & & & & \\ & & \ddots & & & \\ \sigma_1 \tilde{z}_1 & \sigma_2 \tilde{z}_2 & \cdots & \sigma_k \tilde{z}_k & \cdots & \sigma_m \tilde{z}_m \\ & & & & \ddots & \\ & & & & & \sigma_m^2 - \sigma_k^2 \end{pmatrix} \\ &= 2(-1)^{n+1} \sigma_k^{n-m+2} \prod_{i \neq k} (\sigma_i^2 - \sigma_k^2), \end{aligned} \quad (26)$$

where  $\prod_{i \neq k}$  denotes the product for  $i = 1, \dots, k-1, k+1, \dots, m$ . Obviously, from (26),  $J(\alpha_k \mathbf{x}_k)$  is nonsingular.  $\square$

The nonsingularity of  $J(\mathbf{x})$  at  $\mathbf{x} \neq \alpha_k \mathbf{x}_k$  is shown in the following theorem.

**Theorem 4.3.** *The Jacobian matrix  $J(\mathbf{x})$  is nonsingular if and only if*

$$\begin{aligned} &\sigma(\mathbf{v})^{n-m} \left\{ \prod_{i=1}^m (\sigma_i^2 - \sigma(\mathbf{v})^2) \right. \\ &\quad \left. - \frac{1}{C} \sum_{i=1}^m \left[ (\mathbf{u}_i, \mathbf{z}) (\sigma_i^2 (\mathbf{u}_i, \mathbf{u}) + \sigma_i \sigma(\mathbf{v}) (\mathbf{v}_i, \mathbf{v})) \prod_{j \neq i} (\sigma_j^2 - \sigma(\mathbf{v})^2) \right] \right\} \\ &\quad \neq 0. \end{aligned} \quad (27)$$

PROOF. Let us define  $\tilde{M}$ ,  $\tilde{\mathbf{p}}$  and  $\tilde{\mathbf{q}}$  by

$$\tilde{M} = \text{diag}(\sigma_1^2 - \sigma(\mathbf{v})^2, \sigma_2^2 - \sigma(\mathbf{v})^2, \dots, \sigma_m^2 - \sigma(\mathbf{v})^2), \quad (28)$$

$$\tilde{\mathbf{p}} = \begin{pmatrix} \tilde{p}_1 \\ \tilde{p}_2 \\ \vdots \\ \tilde{p}_m \end{pmatrix}, \quad \tilde{p}_i = \sigma_i \tilde{u}_i + \sigma(\mathbf{v}) \tilde{v}_i, \quad (29)$$

$$\tilde{\mathbf{q}} = \begin{pmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ \vdots \\ \tilde{q}_m \end{pmatrix}, \quad \tilde{q}_i = -\sigma_i \tilde{z}_i C^{-1}. \quad (30)$$

Then, from Lemma 4.1,

$$\det(J(\mathbf{x})) = (-1)^n \sigma(\mathbf{v})^{n-m} \det(\tilde{M} + \tilde{\mathbf{p}}\tilde{\mathbf{q}}^T). \quad (31)$$

In the case where  $\tilde{M}$  is nonsingular, it is easily proved (cf. [12]) that

$$\det(\tilde{M} + \tilde{\mathbf{p}}\tilde{\mathbf{q}}^T) = \det(\tilde{M})(1 + (\tilde{\mathbf{q}}, \tilde{M}^{-1}\tilde{\mathbf{p}})).$$

Let  $\text{adj}(\tilde{M})$  denote the adjoint matrix of  $\tilde{M}$ . Since  $\tilde{M}^{-1} = \text{adj}(\tilde{M})/\det(\tilde{M})$ , it follows that

$$\det(\tilde{M} + \tilde{\mathbf{p}}\tilde{\mathbf{q}}^T) = \det(\tilde{M}) + (\tilde{\mathbf{q}}, \text{adj}(\tilde{M})\tilde{\mathbf{p}}). \quad (32)$$

Also, in the case where  $\tilde{M}$  is singular, (32) holds. From (28), it is obvious that

$$\text{adj}(\tilde{M}) = \text{diag}\left(\prod_{j \neq 1}(\sigma_j^2 - \sigma(\mathbf{v})^2), \prod_{j \neq 2}(\sigma_j^2 - \sigma(\mathbf{v})^2), \dots, \prod_{j \neq m}(\sigma_j^2 - \sigma(\mathbf{v})^2)\right). \quad (33)$$

Substituting (28), (29), (30), (32), (33) into (31), we have

$$\begin{aligned} \det(J(\mathbf{x})) &= (-1)^n \sigma(\mathbf{v})^{n-m} \left\{ \prod_{i=1}^m (\sigma_i^2 - \sigma(\mathbf{v})^2) \right. \\ &\quad \left. - \frac{1}{C} \sum_{i=1}^m \left[ \tilde{z}_i (\sigma_i^2 \tilde{u}_i + \sigma_i \sigma(\mathbf{v}) \tilde{v}_i) \prod_{j \neq i} (\sigma_j^2 - \sigma(\mathbf{v})^2) \right] \right\}. \end{aligned}$$

This implies that  $J(\mathbf{x})$  is nonsingular if and only if (27) holds.  $\square$

The case where  $\mathbf{x} = \alpha_k \mathbf{x}_k$  in Theorem 4.3 yields Theorem 4.2 immediately. From Theorems 4.2 and 4.3, we also derive a theorem for the boundedness of the inverse of  $J(\mathbf{x})$ .

**Theorem 4.4.** *If (27) holds, then there exists a positive constant  $M_1$  such that  $\|(J(\mathbf{x}))^{-1}\| \leq M_1$ .*

PROOF. Theorem 4.3 guarantees the existence of  $(J(\mathbf{x}))^{-1}$  and  $|\det(J(\mathbf{x}))| \neq 0$ . Note here that

$$\|(J(\mathbf{x}))^{-1}\| \leq \frac{1}{|\det(J(\mathbf{x}))|} \|\text{adj}(J(\mathbf{x}))\|.$$

Since  $|\det(J(\mathbf{x}))|$  and  $\|\text{adj}(J(\mathbf{x}))\|$  are bounded, we have  $\|(J(\mathbf{x}))^{-1}\| \leq M_1$ .  $\square$

As a result, it is shown that the Jacobian matrix  $J(\mathbf{x})$  is nonsingular and bounded at  $\mathbf{x} = \alpha_k \mathbf{x}_k$  if  $A$  has distinct and nonzero singular values. The Jacobian matrix  $J(\mathbf{x})$  is also generically nonsingular and bounded at  $\mathbf{x} \neq \alpha_k \mathbf{x}_k$ . It is concluded that, in Algorithm 1,  $(J(\mathbf{x}^{(\ell)}))^{-1}$  and  $\mathbf{x}^{(\ell)}$  are generically computable for all  $\ell$ .

## 5. Quadratic convergence

It is well-known [14] that standard Newton's method quadratically converges if the Jacobian matrix is invertible. Algorithm 1 is easily expected to have the local convergence similar to Newton's method, since the algorithm is based on the Newton's method. In this section, we clarify the asymptotic behaviour of Algorithm 1 by regarding the mapping from  $\mathbf{x}^{(\ell)}$  to  $\mathbf{x}^{(\ell+1)}$  as a composition of two mappings.

We first give two lemmas concerning the relations of  $\mathbf{H}(\mathbf{x})$  in (8) and  $\Phi(\mathbf{x})$  in (11) to the Jacobian matrix  $J(\mathbf{x})$ .

**Lemma 5.1.**  *$\mathbf{H}(\mathbf{x})$  satisfies*

$$\mathbf{H}(\mathbf{x} + \Delta \mathbf{x}) = \mathbf{H}(\mathbf{x}) + J(\mathbf{x})\Delta \mathbf{x} - \sigma(\Delta \mathbf{v})\Delta \mathbf{x}, \quad (34)$$

where

$$\Delta \mathbf{x} = \begin{pmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{v} \end{pmatrix}.$$

PROOF. Let us recall that

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} A\mathbf{v} - \mathbf{u}(\mathbf{w}, \mathbf{v})C^{-1} \\ A^T\mathbf{u} - \mathbf{v}(\mathbf{w}, \mathbf{v})C^{-1} \end{pmatrix}.$$

Then  $\mathbf{H}(\mathbf{x} + \Delta\mathbf{x})$  is given by

$$\begin{aligned} \mathbf{H}(\mathbf{x} + \Delta\mathbf{x}) &= \begin{pmatrix} A\mathbf{v} - \mathbf{u}(\mathbf{w}, \mathbf{v})C^{-1} \\ A^T\mathbf{u} - \mathbf{v}(\mathbf{w}, \mathbf{v})C^{-1} \end{pmatrix} \\ &\quad + \begin{pmatrix} -\sigma(\mathbf{v})I_m & A - \mathbf{u}\mathbf{w}^TC^{-1} \\ A^T & -\sigma(\mathbf{v})I_n - \mathbf{v}\mathbf{w}^TC^{-1} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{v} \end{pmatrix} - \frac{(\mathbf{w}, \Delta\mathbf{v})}{C} \begin{pmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{v} \end{pmatrix}. \end{aligned}$$

On the right-hand side, the 1st term is just to  $\mathbf{H}(\mathbf{x})$ . The 2nd includes the Jacobian matrix  $J(\mathbf{x})$  in (12). From the definition of  $\sigma(\cdot)$  in (6), the 3rd becomes  $\sigma(\Delta\mathbf{v})\Delta\mathbf{x}$ . Thus, we have (34).  $\square$

**Lemma 5.2.** As  $\|\Delta\mathbf{x}\| \rightarrow 0$ ,  $\Phi(\mathbf{x} + \Delta\mathbf{x})$  is given by

$$\begin{aligned} \Phi(\mathbf{x} + \Delta\mathbf{x}) &= \Phi(\mathbf{x}) - (J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})(J(\mathbf{x}))^{-1}\mathbf{H}(\mathbf{x}) - \sigma(\Delta\mathbf{v})(J(\mathbf{x}))^{-1}\Delta\mathbf{x} \\ &\quad - [(J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})]^2(J(\mathbf{x}))^{-1}\mathbf{H}(\mathbf{x}) + \boldsymbol{\eta}(\|\Delta\mathbf{x}\|^3), \end{aligned} \quad (35)$$

where

$$\begin{aligned} J'(\Delta\mathbf{x}) &= \begin{pmatrix} \sigma(\Delta\mathbf{v})I_m & \Delta\mathbf{u}\mathbf{w}^TC^{-1} \\ O_{n,m} & \sigma(\Delta\mathbf{v})I_n + \Delta\mathbf{v}\mathbf{w}^TC^{-1} \end{pmatrix}, \quad (36) \\ \boldsymbol{\eta}(\|\Delta\mathbf{x}\|^3) &= \begin{pmatrix} O(\|\Delta\mathbf{x}\|^3) \\ O(\|\Delta\mathbf{x}\|^3) \\ \vdots \\ O(\|\Delta\mathbf{x}\|^3) \end{pmatrix}. \end{aligned}$$

PROOF. From the definition  $\sigma(\cdot)$  in (6) and  $J(\mathbf{x})$  in (12), it follows that

$$J(\mathbf{x} + \Delta\mathbf{x}) = J(\mathbf{x}) - J'(\Delta\mathbf{x}). \quad (37)$$

Let us recall that  $\Phi(\mathbf{x}) = \mathbf{x} - (J(\mathbf{x}))^{-1}\mathbf{H}(\mathbf{x})$ . By taking account of (34) and (37), we derive

$$\begin{aligned} \Phi(\mathbf{x} + \Delta\mathbf{x}) &= \mathbf{x} + \Delta\mathbf{x} \\ &\quad - (J(\mathbf{x}) - J'(\Delta\mathbf{x}))^{-1}(\mathbf{H}(\mathbf{x}) + J(\mathbf{x})\Delta\mathbf{x} - \sigma(\Delta\mathbf{v})\Delta\mathbf{x}). \end{aligned} \quad (38)$$



Note here that  $(J(\mathbf{x}) - J'(\Delta\mathbf{x}))^{-1} = [I_{m+n} - (J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})]^{-1}(J(\mathbf{x}))^{-1}$  in (38). Moreover, by  $[I_{m+n} - (J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})]^{-1} = \sum_{i=0}^{\infty} [(J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})]^i$ , we obtain

$$\begin{aligned}\Phi(\mathbf{x} + \Delta\mathbf{x}) &= \Phi(\mathbf{x}) - (J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})(J(\mathbf{x}))^{-1}\mathbf{H}(\mathbf{x}) \\ &\quad + (J(\mathbf{x}))^{-1}(\sigma(\Delta\mathbf{v})I_{m+n} - J'(\Delta\mathbf{x}))\Delta\mathbf{x} \\ &\quad - [(J(\mathbf{x}))^{-1}J'(\Delta\mathbf{x})]^2(J(\mathbf{x}))^{-1}\mathbf{H}(\mathbf{x}) + \boldsymbol{\eta}(\|\Delta\mathbf{x}\|^3). \quad (39)\end{aligned}$$

Obviously, from (36),  $(\sigma(\Delta\mathbf{v})I_{m+n} - J'(\Delta\mathbf{x}))\Delta\mathbf{x} = -\sigma(\Delta\mathbf{v})\Delta\mathbf{x}$ . By combining it with (39), we have (35).  $\square$

We next show a theorem for the convergence of  $\Phi$  with the help of Lemma 5.2.

**Theorem 5.3 (Quadratic convergence of  $\Phi$ ).** *Let  $\mathbf{z}$  satisfy  $(\mathbf{z}, \mathbf{u}_k) \neq 0$ . Suppose that  $\mathbf{x}^{(0)}$  is sufficiently close to  $\alpha_k\mathbf{x}_k$ . Then  $\mathbf{x}^{(\ell)}$ , generated by (10), quadratically converges to  $\alpha_k\mathbf{x}_k$  as  $\ell \rightarrow \infty$ .*

PROOF. Let  $\mathbf{x}^* = \alpha_k\mathbf{x}_k$  be

$$\mathbf{x}^* = \begin{pmatrix} \mathbf{u}^* \\ \mathbf{v}^* \end{pmatrix}.$$

Let  $\mathbf{x} = \mathbf{x}^*$  and  $\Delta\mathbf{x} = \mathbf{x}^{(\ell)} - \mathbf{x}^*$  in (35). Noting that  $\mathbf{H}(\mathbf{x}^*) = \mathbf{0}$ ,  $\Phi(\mathbf{x}^*) = \mathbf{x}^*$  and  $\mathbf{x}^{(\ell+1)} = \Phi(\mathbf{x}^{(\ell)})$ , it follows that

$$\mathbf{x}^{(\ell+1)} - \mathbf{x}^* = -\sigma(\mathbf{v}^{(\ell)} - \mathbf{v}^*)(J(\mathbf{x}^*))^{-1}(\mathbf{x}^{(\ell)} - \mathbf{x}^*) + \boldsymbol{\eta}(\|\mathbf{x}^{(\ell)} - \mathbf{x}^*\|^3).$$

Here, there exists some positive constant  $M_2$  such that

$$\begin{aligned}|\sigma(\mathbf{v}^{(\ell)} - \mathbf{v}^*)| &= \left| \frac{1}{C} \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix}^T (\mathbf{x}^{(\ell)} - \mathbf{x}^*) \right| \\ &\leq M_2 \|\mathbf{x}^{(\ell)} - \mathbf{x}^*\|. \quad (40)\end{aligned}$$

Since  $\|(J(\mathbf{x}^*))^{-1}\| \leq M_1$ , we have

$$\|\mathbf{x}^{(\ell+1)} - \mathbf{x}^*\| \leq M_1 M_2 \|\mathbf{x}^{(\ell)} - \mathbf{x}^*\|^2. \quad (41)$$

The inequality (41) signifies that  $\mathbf{x}^{(\ell)}$  quadratically converges to  $\mathbf{x}^*$  as  $\ell \rightarrow \infty$ .  $\square$

Finally, we clarify that Algorithm 1 also quadratically converges. As shown in Section 2,  $\mathbf{u}^{(\ell+1)}$  is always on the hyperplane  $\Gamma$  by the mapping  $\Psi$  in (18) after the mapping  $\Phi$  in (11). By analyzing the composition  $\Psi \circ \Phi$  along the line similar to Theorem 5.3, we obtain the following theorem for the convergence of Algorithm 1.

**Theorem 5.4 (Quadratic convergence of Algorithm 1).** *Let  $\mathbf{z}$  satisfy  $(\mathbf{z}, \mathbf{u}_k) \neq 0$ . Suppose that  $\mathbf{x}^{(0)}$  is sufficiently close to  $\alpha_k \mathbf{x}_k$ . Then  $\mathbf{x}^{(\ell)}$ , generated by (17), quadratically converges to  $\alpha_k \mathbf{x}_k$  as  $\ell \rightarrow \infty$ .*

PROOF. Let us here consider a Taylor series of  $\Psi(\mathbf{x} + \Delta \mathbf{x})$ . Since  $(\mathbf{z}, \mathbf{u} + \Delta \mathbf{u}) = (\mathbf{z}, \mathbf{u}) + (\mathbf{z}, \Delta \mathbf{u})$ , we obtain

$$\Psi(\mathbf{x} + \Delta \mathbf{x}) = \frac{C}{(\mathbf{z}, \mathbf{u})} \left( 1 + \frac{(\mathbf{z}, \Delta \mathbf{u})}{(\mathbf{z}, \mathbf{u})} \right)^{-1} (\mathbf{x} + \Delta \mathbf{x}).$$

By combining it with

$$\left( 1 + \frac{(\mathbf{z}, \Delta \mathbf{u})}{(\mathbf{z}, \mathbf{u})} \right)^{-1} = 1 - \left( \frac{(\mathbf{z}, \Delta \mathbf{u})}{(\mathbf{z}, \mathbf{u})} \right) + \left( \frac{(\mathbf{z}, \Delta \mathbf{u})}{(\mathbf{z}, \mathbf{u})} \right)^2 - \dots,$$

we have

$$\begin{aligned} \Psi(\mathbf{x} + \Delta \mathbf{x}) &= \frac{C}{(\mathbf{z}, \mathbf{u})} \mathbf{x} + \frac{C}{(\mathbf{z}, \mathbf{u})} \Delta \mathbf{x} - \frac{C(\mathbf{z}, \Delta \mathbf{u})}{(\mathbf{z}, \mathbf{u})^2} \mathbf{x} \\ &\quad - \frac{C(\mathbf{z}, \Delta \mathbf{u})}{(\mathbf{z}, \mathbf{u})^2} \Delta \mathbf{x} + \frac{C(\mathbf{z}, \Delta \mathbf{u})^2}{(\mathbf{z}, \mathbf{u})^3} \mathbf{x} + \eta(\|\Delta \mathbf{x}\|^3). \end{aligned} \quad (42)$$

Let us here define  $\Delta \mathbf{x}^{(\ell)} = \mathbf{x}^{(\ell)} - \mathbf{x}^*$ . Let  $\mathbf{x} = \mathbf{x}^*$  and  $\Delta \mathbf{x} = \Delta \mathbf{x}^{(\ell)}$  in (35). Then, from  $\mathbf{H}(\mathbf{x}^*) = \mathbf{0}$  and  $\Phi(\mathbf{x}^*) = \mathbf{x}^*$ , it follows that

$$\Phi(\mathbf{x}^{(\ell)}) = \mathbf{x}^* - \sigma(\Delta \mathbf{v}^{(\ell)})(J(\mathbf{x}^*))^{-1} \Delta \mathbf{x}^{(\ell)} + \eta(\|\Delta \mathbf{x}^{(\ell)}\|^3), \quad (43)$$

where

$$\Delta \mathbf{x}^{(\ell)} = \begin{pmatrix} \Delta \mathbf{u}^{(\ell)} \\ \Delta \mathbf{v}^{(\ell)} \end{pmatrix}.$$

Let  $\mathbf{x} = \mathbf{x}^*$  and  $\Delta \mathbf{x} = -\sigma(\Delta \mathbf{v}^{(\ell)})(J(\mathbf{x}^*))^{-1} \Delta \mathbf{x}^{(\ell)} + \eta(\|\Delta \mathbf{x}^{(\ell)}\|^3)$  in (42). Noting that  $(\mathbf{z}, \mathbf{u}^*) = C$  and (43), we obtain

$$\begin{aligned} \Psi(\Phi(\mathbf{x}^{(\ell)})) &= \mathbf{x}^* - \sigma(\Delta \mathbf{v}^{(\ell)}) \left( I_{m+n} - \frac{1}{C} \mathbf{x}^* \begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix}^T \right) (J(\mathbf{x}^*))^{-1} \Delta \mathbf{x}^{(\ell)} \\ &\quad + \eta(\|\Delta \mathbf{x}^{(\ell)}\|^3). \end{aligned} \quad (44)$$

Thus, the difference between  $\mathbf{x}^{(\ell+1)} = \Psi(\Phi(\mathbf{x}^{(\ell)}))$  and  $\mathbf{x}^*$  is evaluated as

$$\begin{aligned} & \|\mathbf{x}^{(\ell+1)} - \mathbf{x}^*\| \\ & \leq \left\| -\sigma(\mathbf{v}^{(\ell)} - \mathbf{v}^*) \left( I_{m+n} - \frac{1}{C} \mathbf{x}^* \begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix}^T \right) (J(\mathbf{x}^*))^{-1} (\mathbf{x}^{(\ell)} - \mathbf{x}^*) \right\|. \end{aligned} \quad (45)$$

Here, there exists some positive constant  $M_3$  such that

$$\left\| I_{m+n} - \frac{1}{C} \mathbf{x}^* \begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix}^T \right\| \leq M_3.$$

Therefore, by using (40) and Theorem 4.4, we have the inequality

$$\|\mathbf{x}^{(\ell+1)} - \mathbf{x}^*\| \leq M_1 M_2 M_3 \|\mathbf{x}^{(\ell)} - \mathbf{x}^*\|^2, \quad (46)$$

which claims that  $\mathbf{x}^{(\ell)}$  quadratically converges to  $\mathbf{x}^*$ .  $\square$

According to Theorem 3.4, if Algorithm 1 converges, then Algorithm 2 also converges. The quadratic convergence of Newton type iteration in Algorithm 1 is shown in Theorem 5.4. In concluding, Algorithm 2 has, so to speak, sequentially the quadratic convergence.

## 6. Numerical examples

In this section, we make sure that Algorithm 1 quadratically converges and SVD is computable by Algorithm 2 through numerical examples. In addition, we show the orthogonality of computed singular vectors. Numerical examples have been carried out on our computer with CPU: Intel Core i7 3.20GHz, Memory: 3GB, OS: Linux kernel 2.6.26 and Compiler: gcc 4.3.2. The parameters in Algorithm 2 are given as  $C = 1$ ,  $\varepsilon_{\text{tol}} = 10^{-13}$ ,  $\varepsilon = 10^{-12}$  and  $\ell_{\text{max}} = 50$ . We compute with double-precision floating point arithmetic. We also adopt CLAPACK 3.0 [15] routine dgesv for computing only  $(J^{(\ell)})^{-1} \mathbf{H}^{(\ell)}$  in the 6th line in Algorithm 1.

The target matrices  $A \in \mathbf{R}^{m \times n}$  of SVD are given by the product of two orthogonal matrices  $U \in \mathbf{R}^{m \times m}$ ,  $V \in \mathbf{R}^{n \times n}$  and  $\Sigma = (\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m) O_{n-m, m}) \in \mathbf{R}^{m \times n}$  in finite arithmetic, namely,  $A = U \Sigma V^T$ . The orthogonal matrices  $U$ ,  $V$  are randomly chosen. It is worth noting that the singular values of  $A$  become  $\sigma_1, \sigma_2, \dots, \sigma_m$ . In other words, we can prepare the target matrices  $A$  with the desired singular values.

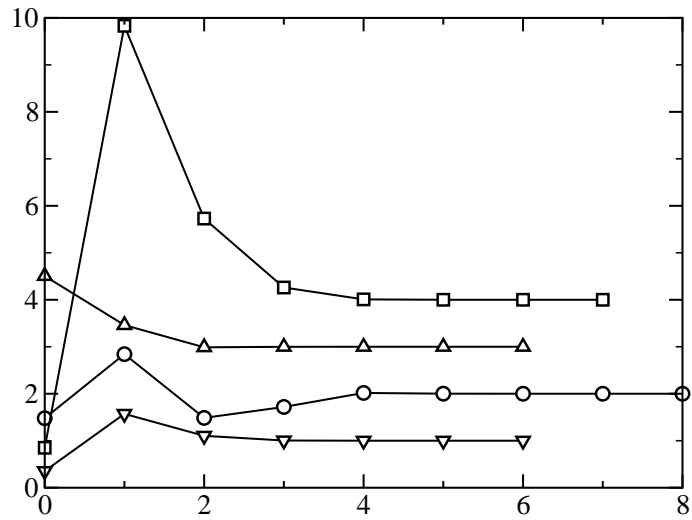


Figure 1: A graph of the iteration number  $\ell$  ( $x$ -axis) and the value of  $\sigma^{(\ell)}$  ( $y$ -axis) in Newton type iteration when all the singular pairs of  $A_1$  are computed by Algorithm 2. The symbols  $\bigcirc$ ,  $\triangle$ ,  $\nabla$  and  $\square$  are in the processes for computing the 1st, the 2nd, the 3rd and the 4th singular pairs, respectively.

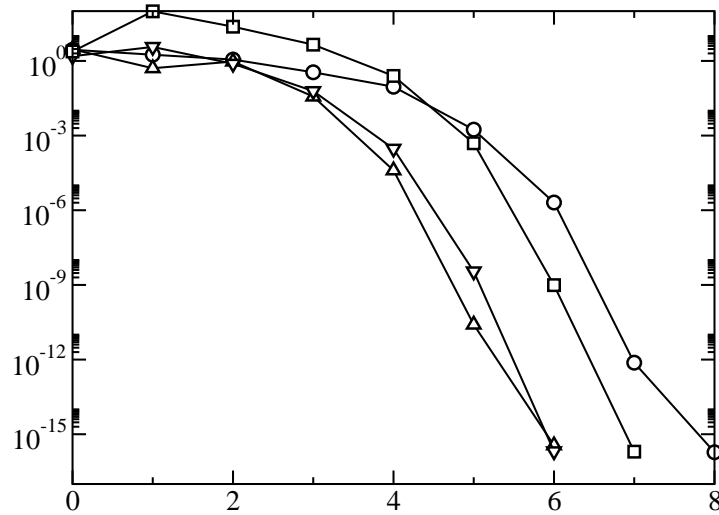


Figure 2: A graph of the iteration number  $\ell$  ( $x$ -axis) and the residual norm  $E^{(\ell)}$  ( $y$ -axis) in Newton type iteration when all singular pairs of  $A_1$  are computed by Algorithm 2. The symbols  $\bigcirc$ ,  $\triangle$ ,  $\nabla$  and  $\square$  are in the processes for computing the 1st, the 2nd, the 3rd and the 4th singular pairs, respectively.

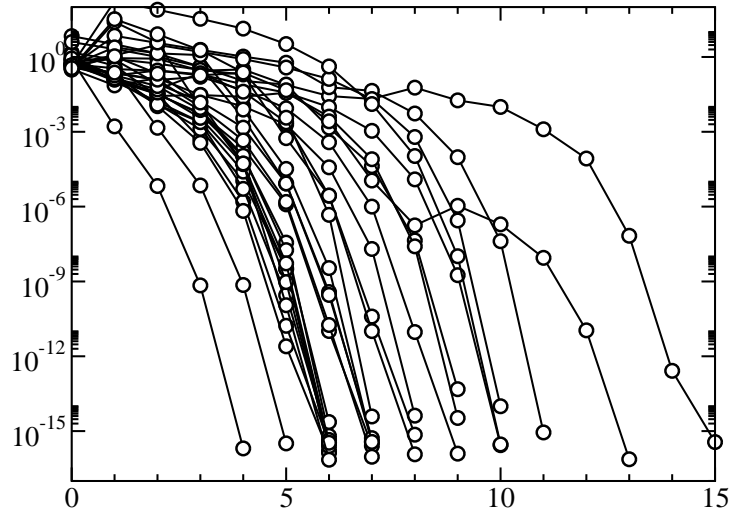


Figure 3: A graph of the iteration number  $\ell$  ( $x$ -axis) and the residual norm  $E^{(\ell)}$  ( $y$ -axis) in Newton type iteration when all singular pairs of  $A_2$  are computed by Algorithm 2. Every plot is in the process for computing the 1st, the 2nd,  $\dots$ , the 30th singular pair.

In first example, we use the matrix  $A_1 \in \mathbf{R}^{4 \times 6}$ , given as the above, with the singular values  $\sigma_1 = 1$ ,  $\sigma_2 = 2$ ,  $\sigma_3 = 3$  and  $\sigma_4 = 4$ . Let us recall that  $\ell$  is the iteration number of Newton type iteration and  $\sigma^{(\ell)}$  in Algorithm 1 converges to one of the singular values as  $\ell \rightarrow \infty$  theoretically. Moreover,  $k$  in Algorithm 2 is regarded as a counter such that the  $k$ th singular pair is being computed by Algorithm 2. It is remarkable here that the  $k$ th singular pair does not always become the pair of the singular value  $\sigma_k$  and the singular vectors corresponding to  $\sigma_k$ . In fact, in this example, the 1st, the 2nd, the 3rd and the 4th singular pairs are the singular pairs related to  $\sigma_2 = 2$ ,  $\sigma_3 = 3$ ,  $\sigma_1 = 1$  and  $\sigma_4 = 4$ , respectively. Figure 1 shows the behaviour of  $\sigma^{(\ell)}$  as the iteration number  $\ell$  grows in the processes for computing the 1st, the 2nd, the 3rd and the 4th singular pairs, respectively. As shown in Figure 1, in the process for the 1st singular pair,  $\sigma^{(\ell)}$ , plotted as  $\bigcirc$ , converges to  $\sigma_2 = 2$  by 8 times Newton type iterations. In the process for the 2nd,  $\sigma^{(\ell)}$ , plotted as  $\triangle$ , converges to  $\sigma_3 = 3$  by 6 times iterations. Similarly, in the process for the 3rd and the 4th,  $\sigma^{(\ell)}$ , plotted as  $\nabla$  and  $\square$ , converge to  $\sigma_1 = 1$  and  $\sigma_4 = 4$  by 6 times and 7 times iterations, respectively. Thus, it turns out that all singular pairs are computed by Algorithm 2. It is remarkable here that the iteration number  $\ell$  has not been beyond  $\ell_{\max} = 50$ , and that Newton type

iteration has not stopped coercively. For the observation of convergence rate, let us introduce the residual norm of  $\mathbf{x}^{(\ell)}$ ,

$$E^{(\ell)} = \frac{\|\mathbf{H}(\mathbf{x}^{(\ell)})\|_{\infty}}{\|\mathbf{u}^{(\ell)}\|_{\infty}}. \quad (47)$$

Figure 2 describes that the residual norm  $E^{(\ell)}$  becomes smaller as the iteration number  $\ell$  increases in the processes for the 1st, the 2nd, the 3rd and the 4th singular pairs, respectively. We see from Figure 2 that Newton type iteration, namely, Algorithm 1 quadratically converges in every process. Algorithm 2, by the way, requires 0.003 seconds for the SVD of  $A_1$ . The orthogonality of computed singular vectors is evaluated as

$$E_{OU} = \max_{i,j=1,2,\dots,m} |(\mathbf{u}_i, \mathbf{u}_j) - \delta_{ij}|, \quad E_{OV} = \max_{i,j=1,2,\dots,m} |(\mathbf{v}_i, \mathbf{v}_j) - \delta_{ij}|. \quad (48)$$

In this example,  $E_{OU} = 1.1 \times 10^{-16}$  and  $E_{OV} = 6.6 \times 10^{-17}$ . This implies that the computed singular vectors have extremely high orthogonality.

In second example, we provide the matrix  $A_2 \in \mathbf{R}^{30 \times 60}$  whose singular values are given by  $\sigma_i = 1 - (1 - 10^{-7})(i - 1)/29$  for  $i = 1, 2, \dots, 30$ . It should be emphasised that the condition number of  $A_2$  is comparatively large as  $10^7$ . Here the condition number is the ratio of the maximal singular value to the minimal one. Figure 3 demonstrates the relationship between the iteration number  $\ell$  and the residual norm  $E^{(\ell)}$  in the processes for the 1st, the 2nd,  $\dots$ , the 30th singular pairs, respectively. Figure 3 suggests that the maximal iteration number of Newton type iteration is 15 and every Newton type iteration quadratically converges. As in first example, Newton type iteration has not failed to compute the singular pair. Of course, all singular pairs are computed by Algorithm 2. It takes 0.137 seconds for the SVD of  $A_2$ . With respect to the orthogonality of the computed singular vectors,  $E_{OU} = 3.8 \times 10^{-13}$  and  $E_{OV} = 3.8 \times 10^{-13}$ . It is also verified that the computed singular vectors have high orthogonality.

## 7. Conclusion

In this paper, we propose a new SVD algorithm based on solving non-linear systems whose solutions are constrained on hyperplanes. To compute each pair of singular value and singular vectors, Newton type iteration is repeatedly employed in our SVD algorithm. The characteristic parameter

in Newton type iteration, namely, normal vector of hyperplane is given from the orthogonal complement of subspace spanned by already obtained singular vectors.

We first show that all pairs are computable if Newton type iteration always converges in our SVD algorithm. We next clarify the nonsingularity and the boundedness of the Jacobian matrix appeared in Newton type iteration. It is shown that the Jacobian matrix is generically nonsingular and bounded if target matrix of SVD has distinct nonzero singular values. Finally, we prove that Newton type iteration quadratically converges by taking account of the nonsingularity and the boundedness of the Jacobian matrix. It is concluded that our SVD algorithm has sequentially quadratic convergence.

We also demonstrate some numerical examples for observing the behaviour of our SVD algorithm. Newton type iteration does not fail to converge, and then its quadratic convergence is verified. Simultaneously, we confirm that computed SVD by our algorithm is with high accuracy.

## References

- [1] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, New York, 2001.
- [2] L. Eldén, *Matrix Methods in Data Mining and Pattern Recognition (Fundamentals of Algorithms 4)*, SIAM, Philadelphia, 2007.
- [3] R. Kakarala, P. O. Ogunbona, Signal analysis using a multiresolution form of the singular value decomposition. *IEEE Trans. Image Processing*, 10:724–735 (2001).
- [4] G. H. Golub, C. F. Van Loan, *Matrix Computations 3rd edn*, The Johns Hopkins University Press, Baltimore, MD, 1996.
- [5] H. Rutishauser, *Lectures on Numerical Mathematics*, Birkhauser, Boston, 1990.
- [6] M. Iwasaki, Y. Nakamura, On the convergence of a solution of the discrete Lotka-Volterra system, *Inverse Problems*, 18:1569–1578 (2002).
- [7] M. Iwasaki, Y. Nakamura, An application of the discrete Lotka-Volterra system with variable step-size to singular value computation, *Inverse Problems*, 20:553–563 (2004).

- [8] G. Peters, J. H. Wilkinson, Inverse iteration, ill-conditioned equations and Newton's method, *SIAM Review*, 21:339–360 (1979).
- [9] M. B. Elgindi, A. Kharab, The quadratic method for computing the eigenpairs of a matrix, *Int. J. Comput. Math.*, 73:517–530 (2000).
- [10] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables (Classics in Applied Mathematics 30)*, SIAM, Philadelphia, 2000
- [11] K. Kondo, S. Yasukouchi, M. Iwasaki, Eigendecomposition algorithms solving sequentially quadratic systems by Newton method, *JSIAM Letters*, 1:40–43 (2009).
- [12] R. Piziak, P. L. Odell, *Matrix Theory : from Generalized Inverses to Jordan Form*, Chapman & Hall/CRC, Boca Ration, 2007.
- [13] R. L. Devaney, *A First Course in Chaotic Dynamical Systems*, Perseus Books, Cambridge, 1992.
- [14] I. K. Argyros, *Computational Theory of Iterative Methods*, Elsevier, Amsterdam, 2007.
- [15] CLAPACK, <http://www.netlib.org/clapack/>